

J. GRUM

RECEIVED

AUG 22 2000

1641

CENTER 1600/2900

ENTERED

## RAW SEQUENCE LISTING

DATE: 08/15/2000

PATENT APPLICATION: US/09/125,031A

TIME: 13:52:15

Input Set : A:\660139.app

Output Set: N:\CRF3\08152000\I125031A.raw

```

3 <110> APPLICANT: LONGACRE-ANDRE, SHIRLEY
4   ROTH, CHARLES
5   BARNWELL, JOHN
6   MENDIS, KAMINI
7   NATO, FARIDABANO
9 <120> TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF
10  PLASMODIUM MSP-1
12 <130> FILE REFERENCE: 0660-0139-0XPCT
14 <140> CURRENT APPLICATION NUMBER: 09/125,031A
15 <141> CURRENT FILING DATE: 1999-03-10
17 <150> PRIOR APPLICATION NUMBER: PCT/FR97/00290
18 <151> PRIOR FILING DATE: 1997-02-14
20 <150> PRIOR APPLICATION NUMBER: FR96/01822
21 <151> PRIOR FILING DATE: 1996-02-14
23 <160> NUMBER OF SEQ ID NOS: 15
25 <170> SOFTWARE: PatentIn Ver. 2.1
27 <210> SEQ ID NO: 1
28 <211> LENGTH: 291
29 <212> TYPE: DNA
30 <213> ORGANISM: Artificial Sequence
32 <220> FEATURE:
33 <223> OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
35 <220> FEATURE:
36 <221> NAME/KEY: CDS
37 <222> LOCATION: (1)..(291)
39 <400> SEQUENCE: 1
40 gaa ttc aac atc tcg cag cac caa tgc gtg aaa aaa caa tgt ccc gag   48
41 Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
42 1 5 10 15
44 aac tct ggc tgt ttc aga cac ttg gac gag aga gag gag tgt aaa tgt   96
45 Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys
46 20 25 30
48 ctg ctg aac tac aaa cag gag ggc gac aag tgc gtg gag aac ccc aac   144
49 Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn
50 35 40 45
52 ccg acc tgt aac gag aac aac ggc ggc tgt gac gca gac gcc aaa tgc   192
53 Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys
54 50 55 60
56 acc gag gag gac tcg ggc agc aac ggc aag aaa atc acg tgt gag tgt   240
57 Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys
58 65 70 75 80
60 acc aaa ccc gac tcg tac ccg ctg ttc gac ggc atc ttc tgc agc taa   288
61 Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser
62 85 90 95
64 taa 291
68 <210> SEQ ID NO: 2
69 <211> LENGTH: 95

```

Input Set : A:\660139.app  
Output Set: N:\CRF3\08152000\I125031A.raw

file://C:\CRF3\Outhold\Vsrl125031A.htm

RAW SEQUENCE LISTING                      DATE: 08/15/2000  
 PATENT APPLICATION: US/09/125,031A              TIME: 13:52:15

Input Set : A:\660139.app  
 Output Set: N:\CRF3\08152000\I125031A.raw

132 acc gag gag gac tcg ggc agc aac ggc aag aaa atc acg tgt gag tgt 240  
 133 Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys  
 134 65 70 75 80  
 136 acc aaa ccc gac tcg tac ccg ctg ttc gac ggc atc ttc tgc agc tcc 288  
 137 Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser Ser  
 138 85 90 95  
 140 tct aac ttc ttg ggc atc tcg ttc ttg ttg atc ctc atg ttg atc ttg 336  
 141 Ser Asn Phe Leu Gly Ile Ser Phe Leu Leu Ile Leu Met Leu Ile Leu  
 142 100 105 110  
 144 tac agc ttc att taa taa 354  
 145 Tyr Ser Phe Ile  
 146 115  
 149 <210> SEQ ID NO: 5  
 150 <211> LENGTH: 116  
 151 <212> TYPE: PRT  
 152 <213> ORGANISM: Artificial Sequence  
 W--> 153 <220> FEATURE:  
 153 <223> OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC  
 155 <400> SEQUENCE: 5  
 156 Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu  
 157 1 5 10 15  
 158 Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys  
 159 20 25 30  
 160 Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn  
 161 35 40 45  
 162 Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys  
 163 50 55 60  
 164 Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys  
 165 65 70 75 80  
 166 Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser Ser  
 167 85 90 95  
 168 Ser Asn Phe Leu Gly Ile Ser Phe Leu Leu Ile Leu Met Leu Ile Leu  
 169 100 105 110  
 170 Tyr Ser Phe Ile  
 171 115  
 175 <210> SEQ ID NO: 6  
 176 <211> LENGTH: 342  
 177 <212> TYPE: DNA  
 178 <213> ORGANISM: Plasmodium falciparum  
 180 <400> SEQUENCE: 6  
 181 aacatttcac aacaccaatg cgtaaaaaaa caatgtccag aaaattcttg atgtttcaga 60  
 182 catttagatg aaagagaaga atgtaaatgt ttattaaatt acaacaaga aggtgataaa 120  
 183 tgtgttgaaa atccaaatcc tacttgtaac gaaaataatg gtggatgtga tgcagatgcc 180  
 184 aaatgtaccg aagaagattc aggtagcaac ggaaagaaaa tcacatgtga atgtactaaa 240  
 185 cctgattctt atccactttt cgatggtatt ttctgcagtt cctctaactt cttaggaata 300  
 186 tcattcttat taatactcat gttaatatata tacagtttca tt 342  
 189 <210> SEQ ID NO: 7  
 190 <211> LENGTH: 387  
 191 <212> TYPE: DNA

RECEIVED  
 AUG 22 2000  
 TECH CENTER 1600/2900

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/125,031A DATE: 08/15/2000  
 TIME: 13:52:15

Input Set : A:\660139.app  
 Output Set: N:\CRF3\08152000\I125031A.raw

RECEIVED

AUG 22 2000

TECH CENTER 1600/2000

```

192 <213> ORGANISM: Plasmodium falciparum
194 <220> FEATURE:
195 <221> NAME/KEY: CDS
196 <222> LOCATION: (1)..(387)
198 <400> SEQUENCE: 7
199 atg aag gcg cta ctc ttt ttg ttc tct ttc att ttt ttc gtt acc aaa 48
200 Met Lys Ala Leu Leu Phe Leu Phe Ser Phe Ile Phe Phe Val Thr Lys
201 1 5 10 15
203 gaa ttc aac atc tcg cag cac caa tgc gtg aaa aaa caa tgt ccc gag 96
204 Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
205 20 25 30
207 gaa ttc aac atc tcg cag cac caa tgc gtg aaa aaa caa tgt ccc gag 144
208 Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
209 35 40 45
211 aac tct ggc tgt ttc aga cac ttg gac gag aga gag gag tgt aaa tgt 192
212 Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys
213 50 55 60
215 ctg ctg aac tac aaa cag gag ggc gac aag tgc gtg gag aac ccc aac 240
216 Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn
217 65 70 75 80
219 ccg acc tgt aac gag aac aac ggc ggc tgt gac gca gac gcc aaa tgc 288
220 Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys
221 85 90 95
223 acc gag gag gac tcg ggc agc aac ggc aag aaa atc acg tgt gag tgt 336
224 Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys
225 100 105 110
227 acc aaa ccc gac tcg tac ccg ctg ttc gac ggc atc ttc tgc agc taa 384
228 Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser
229 115 120 125
231 taa 387
235 <210> SEQ ID NO: 8
236 <211> LENGTH: 127
237 <212> TYPE: PRT
238 <213> ORGANISM: Plasmodium falciparum
240 <400> SEQUENCE: 8
241 Met Lys Ala Leu Leu Phe Leu Phe Ser Phe Ile Phe Phe Val Thr Lys
242 1 5 10 15
243 Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
244 20 25 30
245 Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
246 35 40 45
247 Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys
248 50 55 60
249 Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn
250 65 70 75 80
251 Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys
252 85 90 95
253 Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys
254 100 105 110

```

RAW SEQUENCE LISTING                      DATE: 08/15/2000  
 PATENT APPLICATION:    US/09/125,031A              TIME: 13:52:15

Input Set : A:\660139.app  
 Output Set: N:\CRF3\08152000\I125031A.raw

RECEIVED

AUG 22 2000

TECH CENTER 1600/2000

```

255 Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser
256      115      120      125
260 <210> SEQ ID NO: 9
261 <211> LENGTH: 330
262 <212> TYPE: DNA
263 <213> ORGANISM: Plasmodium falciparum
265 <220> FEATURE:
266 <221> NAME/KEY: CDS
267 <222> LOCATION: (1)..(330)
269 <400> SEQUENCE: 9
270 gaa aca gaa agt tat aag cag ctt gta gcc aac gtg gac gaa ttc aac 48
271 Glu Thr Glu Ser Tyr Lys Gln Leu Val Ala Asn Val Asp Glu Phe Asn
272      1      5      10      15
274 atc tcg cag cac caa tgc gtg aaa aaa caa tgt ccc gag aac tct ggc 96
275 Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu Asn Ser Gly
276      20      25      30
278 tgt ttc aga cac ttg gac gag aga gag gag tgt aaa tgt ctg ctg aac 144
279 Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys Leu Leu Asn
280      35      40      45
282 tac aaa cag gag ggc gac aag tgc gtg gag aac ccc aac ccg acc tgt 192
283 Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn Pro Thr Cys
284      50      55      60
286 aac gag aac aac ggc ggc tgt gac gca gac gcc aaa tgc acc gag gag 240
287 Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys Thr Glu Glu
288      65      70      75      80
290 gac tcg ggc agc aac ggc aag aaa atc acg tgt gag tgt acc aaa ccc 288
291 Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys Thr Lys Pro
292      85      90      95
294 gac tcg tac ccg ctg ttc gac ggc atc ttc tgc agc taa taa 330
295 Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser
W--> 296      100      105      110
299 <210> SEQ ID NO: 10
300 <211> LENGTH: 108
301 <212> TYPE: PRT
302 <213> ORGANISM: Plasmodium falciparum
304 <400> SEQUENCE: 10
305 Glu Thr Glu Ser Tyr Lys Gln Leu Val Ala Asn Val Asp Glu Phe Asn
306      1      5      10      15
307 Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu Asn Ser Gly
308      20      25      30
309 Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys Leu Leu Asn
310      35      40      45
311 Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn Pro Thr Cys
312      50      55      60
313 Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys Thr Glu Glu
314      65      70      75      80
315 Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys Thr Lys Pro
316      85      90      95
317 Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser

```

VERIFICATION SUMMARY  
PATENT APPLICATION: US/09/125,031A

DATE: 08/15/2000  
TIME: 13:52:16

Input Set : A:\660139.app  
Output Set: N:\CRF3\08152000\I125031A.raw

L:72 M:258 W: Mandatory Feature missing, <220> FEATURE:  
L:153 M:258 W: Mandatory Feature missing, <220> FEATURE:  
L:296 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:9

RECEIVED  
AUG 22 2000  
TECH CENTER 1000/2000